



SEQUENCE LISTING

<110> Genencor International, Inc..
Jones, Brian E.
Grant, William D.
Heaphy, Shaun
Grant, Susan

<120> Novel Bacillus 029cel Cellulase

<130> GC796-2-PCT

<140> PCT/US2004/013258

<141> 2004-04-28

<150> US 60/466,831

<151> 2003-04-29

<160> 3

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 3410

<212> DNA

<213> Bacillus sp.

<220>

<221> misc_feature

<222> (1)...(3410)

<223> isolated from environmental sample from Sonachi Lake, Kenya

<400> 1

atcaacacgc	tggaaagtaa	tttcaagggt	aaggccatcg	gttgccgccc	gggtagaaat	60
gtgcggttg	atttcgttga	gcggcgtcgc	cggcgttcca	ccgagggcat	agcgcagcag	120
gttggcgatg	ccaccggtga	ggccttcggg	gccgcctacg	atgttgtgct	cagccgcca	180
tgcgatgtag	ccgtccggct	cgggttcgct	cgcgggggtg	aagaagacaa	tgtcgtcgag	240
ataaaggttg	ccgcttccgc	tctcaacgcc	gccgaggttg	aattggattt	cgcaaattct	300
cgttagggtc	agcacggaat	cgccgacgag	gtcggctatg	ggaatctgaa	tgcgcccata	360
gggttgggta	cgcggaagg	acacgtagg	acccactttg	tcattgggcg	agacgagccg	420
gacaaagatt	tgggtgcgcc	cctgcgagg	gccttgagg	gcgagagaaa	ggtacgtgag	480
ggcgctgatg	tcgtgcgtg	gaccgtctcc	ccagttgtcg	agattgagcc	caaatccggc	540
ccaccatccg	gcgatagtgt	agctccaatg	gtagtgcgc	tcaccctcga	agccgcccgt	600
ggagagttcc	tgcaagccgt	cgccccaaat	gcccgtgatg	agcgttgcc	cgtaacggta	660
gatacaaatg	tcggcgccg	gtgccgggg	aagatcgcc	tgagtgatca	cgagagtggc	720
gggtggcgctg	ccttcgtgat	tagggctcgt	aatgggtggc	acgaccgtgt	agctaccggg	780
ccccactggc	gcatgggtg	aaccgttgta	ggtaaaggag	acgtcaagcc	ccacgggatg	840
gggtctcgga	agagcgccct	tgggggtgcc	gtcgaaaacg	tgttccaaat	tgagagcgt	900
gatgggtggc	ggtgccttga	gcacagtcac	agaaacagtg	gattgcacgg	gatcgtgcgc	960
tgccgtgtct	gcaggtgtga	agaccacgct	gtaaaaaacg	gttccggcgg	acggtgcaag	1020
gccggacagg	acaaaggcaa	agtcgccggg	gacggcggct	actccgccgc	tcaggccggc	1080
ctccgcaagg	gtttgcccga	aggtgatggg	tgccggtgtg	ggccacatct	ccacaaggcc	1140
gggtgtcccc	tcgtcacgca	ccggcatgag	ggcggagagg	agatgaatgt	aactggcttg	1200
gtaattgatg	tcgggctcgg	tgatttccca	tgagttctcc	ggccaaaaac	cattccaatc	1260
aaggtaggct	ttttgcacgg	gttgggtctc	gatcgccctga	atgcttccgc	tgtatttggg	1320

cattgggacc	cgcccgaag	aaaaccagga	gcgggaccgt	agagtgaagt	gagggcattg	1380
tcccagtcg	gccatcgcg	aaccaatggt	ggtagatttc	attggctgca	cggtcagcgc	1440
cgctggcata	catgttgcta	agatagacca	tgcccatttg	gttactccg	tgagatagat	1500
gcaggtagcc	catcgcgga	tcgcatgcg	cgcccgcgct	ggcggggttg	agcccaagcc	1560
tccgtacccc	ctcgaagaaa	aagccagcct	gagactttgt	tttggttcgag	ccccacgtgt	1620
aatcctgatc	cttcaggtag	gcgcggtagg	cgtcggtctg	gttattccat	gcaccgagaa	1680
actccccacc	gtttatagaa	gccgccatcc	ggttgcggtat	gtcggcagag	acgctaggcg	1740
tcgctcccg	gagggtcgtg	tagtgggcga	gagctttttg	tagctcacct	tgaaagggga	1800
agaaatacca	ccactgcacg	ggctccatat	cgagatagcg	cacatcgaag	aaatcgcgat	1860
agaccgcacc	gcccgtgcgc	tcgaagagca	tggcgcgggc	catcacacgg	ttggctagcg	1920
tatcggtggc	attgcgcgag	gggtcacg	aagcaaatcc	ggtgttgctg	aaaggcacat	1980
gaggatggac	catggtccaa	ttccatgcgg	cgatggcagc	ggattcgagg	gtgacggcat	2040
aatcgctcat	gcctacgctc	tcaaagacag	tcgccccgag	ggcgaaagcg	gcggcagcca	2100
tggcagtg	ctcggtcgag	acggggccgt	agtaacgcgg	atgggtgtcg	gtgctcgcg	2160
ggctggcgct	ctgggtcccc	gtcacggaaa	ctttcccgag	aatagccccg	ctcggtcct	2220
gcatgcgtaa	gagccagtc	attccccatt	tgacttcgtc	aagcaggtcg	gggacaccgt	2280
tgccggtatc	cgggatgcca	aaatcatcgg	taaagacgtc	aggccgcct	tgataggcaa	2340
ggagcagctc	caggatgacg	cgccccgtcc	actcgctgta	cttggtgaaa	tcgcccgc	2400
cgaaccaacc	gccgctgaga	tcgctgcca	aggaggcatt	ccccatatcc	cagatggggc	2460
ggctggcgac	gtcctgcggg	tgagaagcgg	catcgcccc	gttcgctgg	gcgtagggca	2520
cctccttg	aaacccggag	cgctgataga	agaacatgcg	cacggcctcg	cgcaggacaa	2580
catcgtaaac	atccgcgcca	atggcgaaac	tatcggaatg	agtgttggtg	gcaggatcgt	2640
ggatgcggta	gtggccgggc	tcggcaacta	ccgtaaaatc	aaaccaccac	acgcggtctc	2700
ccgattgaat	atggatggcg	ccgccgttcc	acgggaccgg	tgagccggag	aaaaccacga	2760
cgccatcg	cacgcgacgg	acctccagcg	ttgcgcggg	gctgtagctc	tcggcgctgt	2820
tccagccaat	ctgcgggtcg	gcgatcaccg	ccaccttggt	ggcatcgcg	gggtaaccga	2880
attggtcgat	gcggatttta	tcggtgtggg	tggaggcgac	gagggcgagg	ctgcccata	2940
gcagcaagaa	aaagcccgc	gtcggccccg	taccaaaaa	acgaatagg	agagaaaaat	3000
tcatagcagg	atgtggatac	ggaaagggg	aaaacggtgc	aaagacccaa	gcccacgc	3060
tggcgaaaac	tggtatgggtg	gtttatcaag	aaaagcgctt	ttgagccaaa	agctgcgggc	3120
aatccttatt	gcgtttcaca	atattttcac	atcgctcgcg	gcacgacttt	tcgatgggcg	3180
acttgacagc	gtattctctc	aggcgcgagg	ctgcaaacct	tatgaaaaaa	ggcccgcgca	3240
gcgatctgtc	cccgtcaaaa	atccagtc	ggtttgttca	agggtttgag	gtctgataga	3300
ggcacagtcg	agccatcagc	agtcgcattg	agtaggggtg	ttggagaaag	tgtgcaaatg	3360
accgctgccg	aaggaactgt	ggagacaaaa	agcatatttt	cctcgccaag		3410

<210> 2

<211> 1746

<212> DNA

<213> Bacillus sp.

<220>

<221> misc_feature

<222> (1)...(1746)

<223> isolated from environmental sample from Sonachi Lake, Kenya

<400> 2

atgaattttt	ctctccctat	tcgttttttt	ggtatcgggc	cgacagcggg	ctttttcttg	60
ctgctcatgg	gcagctccgc	cctcgctgcc	tccaccacaca	ccgataaaat	ccgcatcgac	120
caattcgggt	accccgccga	tgccaccaag	gtggcggtga	tcgcccagccc	gcagattggc	180
tggaacagcg	ccgagagcta	cagccccggc	gcaacgctgg	aggctccgtcg	cgtgaacgat	240
ggcgctcg	ttttctccgg	ctcaccggtc	ccgtggaacg	gcggcgccat	ccatattcaa	300
tcgggagacc	gcgtgtgg	gtttgatttt	acggtagttg	ccgagcccgg	ccactaccgc	360
atccacgatc	ctgccaacaa	cactcattcc	gatagtttcg	ccattggcgc	ggatgtttac	420
gatgtgtgcc	tgcgcgaggc	cgtgcgc	ttcttctatc	agcgctccgg	gtttgccaag	480
gaggtgccct	acgcccacgc	gaactggggc	gatgccgctt	ctcaccgcga	ggacgtcgcc	540
agccgcccc	tctgggatat	ggggaatgcc	tccttgagagc	gcgatctcag	cggcggttg	600

ttc gat g c g g	g c g a t t t c a a	c a a g t a c a g c	g a g t g g a c g g	g g c g c g t c a t	c c t g g a g c t g	660
c t c c t t g c c t	a t c a a g g g g c g	g c c t g a c g t c	t t t a c c g a t g	a t t t t g g c a t	c c c g g a a t c c	720
g g c a a c g g t g	t c c c c g a c c t	g c t t g a c g a a	g t c a a a t g g g	g a a t g g a c t g	g c t c t t a c g c	780
a t g c a g g a g c	c g a g c g g g g c	t a t t c t c g g g	a a a g t t t c c g	t g a c g g g g c a	c c a g a g c g c c	840
a g c c c g c c g a	g c a c c g a c a c	c c a t c c g c g t	t a c t a c g g c c	c c g t c t c g a c	c g a g g c c a c t	900
g c c a t g g c t g	c c g c c g c t t t	c g c c c t c g g g	g c g a c t g t c t	t t g a g a g c g t	a g g c a t g a g c	960
g a t t a t g c c g	t c a c c c t c g a	a t c c g c t g c c	a t c g c c g c a t	g g a a t t g g a c	c a t g g t c c a t	1020
c c t c a t g t g c	c t t t c g a c a a	c a c c g g a t t t	g c t t c c g t g a	g c c c c t c g c g	c a a t g c c c a c	1080
g a t a c g c t a g	c c a a c c g t g t	g a t g g c c g c c	g c c a t g c t c t	t c g a g c g c a c	g g g c g g t g c g	1140
g t c t a t c g c g	a t t t c t t c g a	t g t g c g c t a t	c t c g a t a t g g	a g c c c g t g c a	g t g g t g g t a t	1200
t t c t t c c c c t	t t c a a g g t g a	g c t a c a a a a a	g c t c t c g c c c	a c t a c a c g a c	c c t c c c g g g a	1260
g c g a c g c c t a	g c g t c t c t g c	c g a c a t c c g c	a a c c g g a t g g	c g g c t t c t a t	a a a c g g t g g g	1320
g a g t t t c t c g	g t g c a t g g a a	t a a c c a g a c c	g a c g c c t a c c	g c g c c t a c c t	g a a g g a t c a g	1380
g a t t a c a c g t	g g g g c t c g a a	c a a a a c a a a g	t c t c a g g c t g	g c t t t t t c t t	c g a g g g g g t a	1440
c g g a g g c t t g	g g g t c a a c c c	c g c c g a c g c g	g c c g c g c a t c	g c g a t g c c g c	g a t g g g c t a c	1500
c t g c a c t a t c	t c c a c g g a g t	g a a c c c a a t g	g g c a t g g t c t	a t c t t a g c a a	c a t g t a t g c c	1560
a g c g g c g c t g	a c c g t g c a g c	c a a t g a a a t c	t a c c a c c a t t	g g t t c c g c g a	t g g c c g g a c t	1620
g g g a c a a t g c	c c t c a c t t c a	c t c t a c g g t c	c c g t c c t g g	t t t t c t t t c g	g g c g g g t c c c	1680
a a t g c c c a a a	t a c a g c g g a a	g c a t t c a g g c	g a t c c g a g a c	c a a c c c g t g c	a a a a a g c c t a	1740
c c t t g a						1746

<210> 3

<211> 581

<212> PRT

<213> Bacillus sp.

<220>

<221> VARIANT

<222> (1)...(581)

<223> isolated from environmental sample from Sonachi Lake, Kenya

<400> 3

Met	Asn	Phe	Ser	Leu	Pro	Ile	Arg	Phe	Phe	Gly	Ile	Gly	Pro	Thr	Ala	
1				5					10					15		
Gly	Phe	Phe	Leu	Leu	Leu	Met	Gly	Ser	Ser	Ala	Leu	Val	Ala	Ser	Thr	
			20					25					30			
His	Thr	Asp	Lys	Ile	Arg	Ile	Asp	Gln	Phe	Gly	Tyr	Pro	Ala	Asp	Ala	
		35				40						45				
Thr	Lys	Val	Ala	Val	Ile	Ala	Asp	Pro	Gln	Ile	Gly	Trp	Asn	Ser	Ala	
	50				55						60					
Glu	Ser	Tyr	Ser	Pro	Gly	Ala	Thr	Leu	Glu	Val	Arg	Arg	Val	Asn	Asp	
65				70						75				80		
Gly	Val	Val	Val	Phe	Ser	Gly	Ser	Pro	Val	Pro	Trp	Asn	Gly	Gly	Ala	
			85					90						95		
Ile	His	Ile	Gln	Ser	Gly	Asp	Arg	Val	Trp	Trp	Phe	Asp	Phe	Thr	Val	
		100						105					110			
Val	Ala	Glu	Pro	Gly	His	Tyr	Arg	Ile	His	Asp	Pro	Ala	Asn	Asn	Thr	
	115					120					125					
His	Ser	Asp	Ser	Phe	Ala	Ile	Gly	Ala	Asp	Val	Tyr	Asp	Val	Val	Leu	
	130					135					140					
Arg	Glu	Ala	Val	Arg	Met	Phe	Phe	Tyr	Gln	Arg	Ser	Gly	Phe	Ala	Lys	
145				150						155				160		
Glu	Val	Pro	Tyr	Ala	His	Ala	Asn	Trp	Ala	Asp	Ala	Ala	Ser	His	Pro	
			165					170					175			
Gln	Asp	Val	Ala	Ser	Arg	Pro	Ile	Trp	Asp	Met	Gly	Asn	Ala	Ser	Leu	
	180							185					190			
Glu	Arg	Asp	Leu	Ser	Gly	Gly	Trp	Phe	Asp	Ala	Gly	Asp	Phe	Asn	Lys	

		195					200					205				
Tyr	Ser	Glu	Trp	Thr	Gly	Arg	Val	Ile	Leu	Glu	Leu	Leu	Leu	Ala	Tyr	
	210					215					220					
Gln	Gly	Arg	Pro	Asp	Val	Phe	Thr	Asp	Asp	Phe	Gly	Ile	Pro	Glu	Ser	
225					230					235					240	
Gly	Asn	Gly	Val	Pro	Asp	Leu	Leu	Asp	Glu	Val	Lys	Trp	Gly	Met	Asp	
				245					250					255		
Trp	Leu	Leu	Arg	Met	Gln	Glu	Pro	Ser	Gly	Ala	Ile	Leu	Gly	Lys	Val	
			260					265					270			
Ser	Val	Thr	Gly	His	Gln	Ser	Ala	Ser	Pro	Pro	Ser	Thr	Asp	Thr	His	
		275					280					285				
Pro	Arg	Tyr	Tyr	Gly	Pro	Val	Ser	Thr	Glu	Ala	Thr	Ala	Met	Ala	Ala	
	290					295					300					
Ala	Ala	Phe	Ala	Leu	Gly	Ala	Thr	Val	Phe	Glu	Ser	Val	Gly	Met	Ser	
305					310					315					320	
Asp	Tyr	Ala	Val	Thr	Leu	Glu	Ser	Ala	Ala	Ile	Ala	Ala	Trp	Asn	Trp	
				325					330					335		
Thr	Met	Val	His	Pro	His	Val	Pro	Phe	Asp	Asn	Thr	Gly	Phe	Ala	Ser	
			340					345					350			
Val	Ser	Pro	Ser	Arg	Asn	Ala	His	Asp	Thr	Leu	Ala	Asn	Arg	Val	Met	
		355					360					365				
Ala	Ala	Ala	Met	Leu	Phe	Glu	Arg	Thr	Gly	Gly	Ala	Val	Tyr	Arg	Asp	
	370					375					380					
Phe	Phe	Asp	Val	Arg	Tyr	Leu	Asp	Met	Glu	Pro	Val	Gln	Trp	Trp	Tyr	
385					390					395					400	
Phe	Phe	Pro	Phe	Gln	Gly	Glu	Leu	Gln	Lys	Ala	Leu	Ala	His	Tyr	Thr	
				405					410					415		
Thr	Leu	Pro	Gly	Ala	Thr	Pro	Ser	Val	Ser	Ala	Asp	Ile	Arg	Asn	Arg	
			420					425					430			
Met	Ala	Ala	Ser	Ile	Asn	Gly	Gly	Glu	Phe	Leu	Gly	Ala	Trp	Asn	Asn	
		435					440					445				
Gln	Thr	Asp	Ala	Tyr	Arg	Ala	Tyr	Leu	Lys	Asp	Gln	Asp	Tyr	Thr	Trp	
	450					455					460					
Gly	Ser	Asn	Lys	Thr	Lys	Ser	Gln	Ala	Gly	Phe	Phe	Phe	Glu	Gly	Val	
465					470					475					480	
Arg	Arg	Leu	Gly	Leu	Asn	Pro	Ala	Asp	Ala	Ala	Ala	His	Arg	Asp	Ala	
				485					490					495		
Ala	Met	Gly	Tyr	Leu	His	Tyr	Leu	His	Gly	Val	Asn	Pro	Met	Gly	Met	
			500					505					510			
Val	Tyr	Leu	Ser	Asn	Met	Tyr	Ala	Ser	Gly	Ala	Asp	Arg	Ala	Ala	Asn	
		515					520					525				
Glu	Ile	Tyr	His	His	Trp	Phe	Arg	Asp	Gly	Arg	Thr	Gly	Thr	Met	Pro	
	530					535					540					
Ser	Leu	His	Ser	Thr	Val	Pro	Leu	Leu	Val	Phe	Phe	Arg	Ala	Gly	Pro	
545					550											